Climate extremes: Data issues

Lisa Alexander
Climate Change Research Centre and ARC Centre of Excellence for Climate Systems Science, University of New South Wales

WCRP-ICTP 21st July 2014
WCRP Grand Challenge on Extremes

• 8 key scientific questions have been posed to the scientific community over the coming decade

Q1: How can we improve the collation, dissemination and quality of observations needed to assess extremes and what new observations do we need?
Why observe the climate?
How do we observe the climate?

This is not the lab!

Michael de Podesta’s ‘Instrument of real beauty’ for determining the Boltzmann constant
Observations in some really challenging environments ...
Many different observations
History of instrumentation and observations

• Between 16th and 17th C instruments started to be constructed to measure weather variables e.g. temperature (Galileo, Santorio), pressure (Torricelli)

• 18th and 19th century
  - 1724 - Farenheit introduces first standardised scale
  - 1742 - Celcius introduces scale based on freezing and boiling point of water
  - 1848 – Kelvin introduces scale based on ‘absolute temperature’

• Military realises benefit of weather observations (first daily forecast produced - 1860)
20th century

Late 1800s/Early 1900s – National Weather Services formed
1930 – first weather balloon released
1941 – first radar network
1959 – first weather satellite launched (Vanguard 2)
1975 – first GOES satellite launched
Data requirements for extremes

- Quality, consistency and availability of observations underpin analysis of climate extremes and attribution analyses

- Errors in data are likely to show up as ‘extreme’

- But what we want/need versus what we get can be very different
Quality – what we want
Quality – what we get
Consistency – what we want

Artificial changes are well documented (metadata) and have been accounted for over all time periods (traceability)
Consistency – what we get

Base anomaly series and regression fit
Availability – what we want

www.get_all_data_from_all_stations_for_all_time_periods_in_my_favorite_format.com
Availability – what we get

Sub-daily precip stations (HadISD) and SDII coverage (HadEX2)

“We estimate the world loses 500,000 old records each day”. Rick Crouthamel, IEDRO
Issues and uncertainties – 1. station error

uncertainty of individual station measurements
Issues and uncertainties: 2. sampling error

uncertainty in area mean caused by sampling small number of point values

Source: Donat et al. 2013
Issues and uncertainties: 3. bias error

...also coverage error...

uncertainty in large-scale values caused by systematic changes in measurement methods
Global SST measurements

But what about 1945?

Source: Thompson et al. 2008
No evidence of change over land

Source: Thompson et al. 2008
Mostly US ships and buoys before 1945, mostly UK after

Source: Thompson et al. 2008
Inhomogeneities

Artificial changes in the observed record that are not due to changes in climate

Can be introduced through changes in e.g.
- observation source or observing practice
- number of observations
- station relocation
- instrumentation change
- site environment

Inhomogeneities can be sudden or gradual
Statistical tests can adjust data for inhomogeneities

\[ T_{obs} + C_H = T_{true} + \varepsilon_{obs} + \varepsilon_H \]

Homogenisation correction

Error in applied homogenisation correction

Random measurement error
Is this different for extremes?

Before a site move in 1993 an ‘extreme’ Tmin < 2°C occurred only 5 times in Perth but almost every year since then.
Observing practice – it never rains on a Sunday

Source: King et al. 2012
The long list of issues to be considered

- Station moves
- Instrument changes
- Observer changes
- Automation
- Time of observation biases
- Microclimate exposure changes
- Urbanization
- And so on ...
- And that’s just for the land records, similar laundry lists for satellites, radiosondes, marine …
Underlying which are two absolutely fundamental issues ...

- A lack of traceability to absolute or even relative standards for most of the historical (and present day) records
- A lack of adequate documentation of the (ubiquitous) changes sufficient to characterize in an absolute sense the time-varying measurement characteristics.
All data have problems and uncertainties

• Which are different from (climate) models
• Which require expertise to minimize
  – E.g. not all data sets were created equal
• Fortunately, there are experts in almost all types of observational data who would be happy to share their insights
  – As these data are often their life’s work
Very likely decreases in cold days and nights and increases in warm days and nights

Likely more land regions where the number of heavy precipitation events has increased than where it has decreased

Low confidence in long-term increases in intense tropical cyclone activity

Virtually certain North Atlantic since 1970s

Low confidence in increasing drought duration/intensity

Likely increases and decreases in some regions
Recent attempts to reconcile differences

Source: Trenberth et al. 2014
Scaling issues
Fundamental mismatch?

• The spatial representativeness of in situ observations which are gridded using interpolation techniques may not ‘mean’ the same as climate model output of extremes

• Scale mismatch (or ‘problem of change in support’) more importantly affects phenomena whose spatial features are discontinuous or have short temporal scales
  – e.g. sub-daily precipitation, extreme events

• Alternative data sets are available (e.g. reanalyses) but come with their own problems
Points versus grids

1. Observations are taken at points locations
2. Climate model output represents an areal average

How can we compare 1. and 2.?
Gridding allows comparison with climate models

Source: Alexander and Arblaster, 2009
Gridding versus downscaling

- Gridding methods usually work by assigning values to unknown points by means of a weighted average of a number of known points.

- Downscaling is the means of converting areally averaged data to point data. There are two main types:
  - Statistical
  - Dynamical
Gridding methods

- There are many techniques that convert point observations to areal averages (grids). These include:
  - Kriging
  - Natural neighbours
  - Minimum curvature
  - Inverse distance weighting
  - Angular distance weighting
  - Radial Basis Functions
  - …
Example: Angular Distance Weighting (ADW)

- Requires knowledge of the correlation structure of the data i.e. how the chosen climate variable varies across space and time

- A correlation function, $f_i$, is defined for each $i^{th}$ station:

\[ f_i = e^{-\frac{r_i}{L}} \]

- Distance from $i^{th}$ station to gridbox centre
- Decorrelation length scale
Decorrelation length scale

- Also referred to as the *e-folding distance*
  - i.e. the time it takes the fitted function to fall away with distance to 1/e
Calculating weighting functions

\[ \omega_i = f_i^m \left\{ \frac{\sum_k f_k^m \left[ 1 - \cos(\theta_k - \theta_i) \right]}{\sum_k f_k^m} \right\}, \quad i \neq k \]

Adjusts function decay

Bearing of \(i^{th}\) and \(k^{th}\) stations

\(k\) sums over all stations within circle of influence (R)

R is determined by the correlation function (usually the e-folding distance)
Testing values of $m$

- The parameter that determines the exponential decay of station weights
Parameter choices

• Period over which correlation is performed (all years, subset of years)
• Value of m (multiple tests can determine ‘best’ value)
• Distance over which function is fitted
• Ultimately parameter choices necessarily have to reflect compromises (don’t have data points for all space and time periods, don’t have all information)
Testing the ‘goodness of fit’ of the chosen method

• The method can be tested by trying to simulate the station data rather than an unknown grid point location
Error fields

• Possible to determine regions where you have more/less confidence in the gridded product

• Enables ‘error bars’ to be placed on gridded observational data

• Important for climate model evaluation

Source: Caesar et al. (2006), JGR
Parametric versus structural uncertainties

• **Parametric uncertainties** can be calculated by testing our choices within our model framework e.g.
  
  ‒ Parameter settings
  ‒ Minimum number of ‘points’ to consider
  ‒ Type of function to fit and maximum distance over which to fit function
  ‒ Value of function decay

• **Structural uncertainties** can be calculated by using multiple gridding methods
  
  ‒ Questioning your methodological framework
  ‒ E.g. Kriging versus ADW

• Need to understand uncertainties before appropriate evaluation of climate models
Comparison of the different data sets

- **HQ stations** (1° grid, fixed parameters)
- **ALL stations**
- **AWAP**

**Hottest day of the year**

- Same gridding method
  - Different input data
- Same input data
  - Different gridding method

---

ARC CENTRE OF EXCELLENCE FOR CLIMATE SYSTEM SCIENCE

UNSW Climate Change Research Centre
Also does grid size matter...?
How large is this uncertainty?

Example for Australia

Several hundred realizations by varying:

- Grid size
- Parameters
- Network density
- Interpolation method
- Prior painstaking work on quality control and homogenisation of data

Source: Vogel, 2012
Can test multiple methods and regions

Source: Yin et al. 2014
Bias estimation
Trend estimation
For absolute extremes quality particularly important
Why don’t we just use reanalysis data?

- Reanalyses based on observational data assimilated into a NWP model
  - ✓ provide physically consistent fields
  - ✓ complete spatial coverage
  - ✗ Usually shorter than in situ records
  - ✗ Inhomogeneous especially since assimilation of satellite records
Advantages and disadvantages of reanalyses

- Able to assess fields for which there are little or no observations
- Output may be more similar to GCMs
- Disagreement between products and between products and observations
- Inhomogeneous over time (especially due to inclusion of satellite data in late 1970s)
Types of reanalysis products

- European Centre for Medium-range Weather Forecasts (ECMWF) e.g.
  - ERA-40 (Sep 1957-Aug 2002)
  - ERA-interim (1979-present)

- USA e.g.
  - NCEP/NCAR reanalysis (1948/01/01 to present)
  - NASA MERRA (1979-present)

- Japan e.g.
  - JRA-25 (1979-2004)
Reanalysis data in comparison to gridded *in situ* data sets

Source: Donat et al. (2014)
Example spatial correlation between datasets

Correlations $\text{TX}_x$

(b) $\text{Rx}_{5\text{day}}$

common years 1979-2010

common years 1979-2010
Example temporal correlation between datasets

Pattern correlation TXx

1.0
0.8
0.6
0.4
0.2

1980  1990  2000  2010

GHCNDEX
HadGHCND
NCEP1
ERA40
NCEP2
ERA-Interim
JRA25
Calculating uncertainties for global datasets of extremes

Structural uncertainties are larger than parametric uncertainties
Some indices have larger uncertainties than others

Source: Dunn et al. (2014)
‘Observations’ versus models

Avg annual precipitation intensity (mm/d) over land (masked)
Remember that observations are models

• “All models are wrong…”
• “..but SOME are useful”

• “All models using erroneous data are wrong…”
• “…and NONE are useful”

G.E.P. Box

L.V. Alexander